(51) International Patent Classification ^{6} :		(11) International Publication Number:	WO 98/36362
G06F 13/42, 15/173	A1	(43) International Publication Date:	20 August 1998 (20.08.98)

(54) Title: INTERNET UPSTREAM REQUEST COMPRESSION
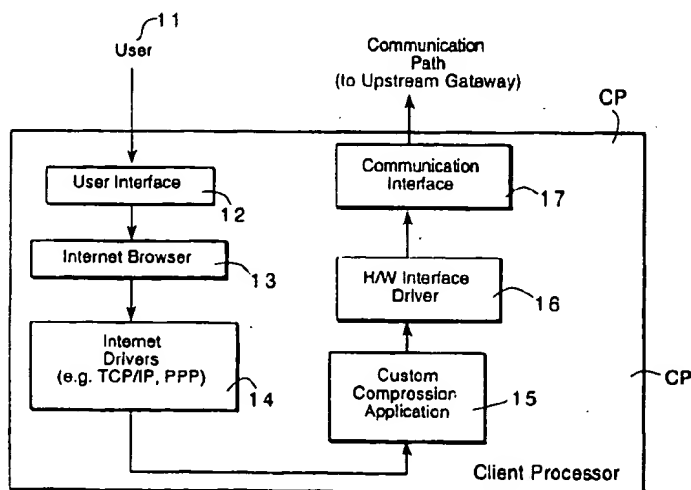
(57) Abstract

Interactive internet (5) activities are a very popular means for gathering information for business, personal, medical, entertainment and other purposes. Most internet interaction is asymmetrical in nature in that a client's (1) requests for information are much smaller (in data size) than the resulting information delivered by the server (7). Although the client/internet link (2, 11) is inherently asymmetrical, there remains a great deal of interaction and overhead required between the client (1) and the server (7) that increases the bandwidth needs for the client's upstream request channel (2, 3, 4). There are many advantages to reducing this upstream traffic volume and thereby making the internet interaction even more asymmetrical. Such compression (15) of request data could allow more clients (1, CP) to utilize a single upstream data path instead of separate paths.

INTERNET UPSTREAM REQUEST COMPRESSION

TECHNICAL FIELD

This invention relates to interactive information requests over the internet.

BACKGROUND ART

5   To gain access to the internet, a client (user) will typically login into an Internet Service Provider (ISP) via a telephone or ISDN modem. This connection is typically a medium data rate (i.e., 9.6 to 28.8 kbps), symmetrical connection. However, the client is typically searching for

10   information, which leads to asymmetric communications. For example, a client (user) requesting a file download will send a small upstream request (e.g., 200 bytes) but receive a large file (e.g. 200 kbytes) in return from the server. In these types of connection the downstream channel (i.e.,

15   return link) is the bottleneck.

The typical volume of downstream data to upstream data ratio is from 10:1 to 20:1. Currently, asymmetrical channel services are being deployed (e.g., satellite broadcast, cable modems, ADSL, etc.) to take advantage of

20   this ratio and reduce this downstream bottleneck. These services typically increase the downstream channel capacity

2

by using a media with wider bandwidth (e.g., co-axial cable), and maintain or allocate a smaller bandwidth connection for the upstream user request channel. In some cases (such as satellite broadcast), an upstream channel

5   path, independent of the downstream channel path, is utilized.

The Internet is a packet switch network, where the defacto protocol standard is TCP/IP. The Internet Protocol (IP) provides the basic addressing scheme for internet

10   routing (where the information goes). Transmission Control Protocol (TCP) is very robust protocol designed for error-free bulk data transfer and error detection and correction. On top of TCP/IP, an additional serial interface, protocol (e.g., PPP or SLIP) is used for connection with an ISP.

15   From the standpoint of upstream channel efficiency, a typical internet request has a high degree of packet overhead to accommodate the layers of protocol, the error correction, and the data coding scheme.

Browsing on the World Wide Web (WWW) (one of the most

20   popular activities on the internet) also leads to inefficient use of the upstream channel. This browsing is based on the HyperText Transfer Protocol (HTTP). Typically, a user sends an HTTP request for a Web Page to a particular server on the internet and the server responds

25   with requested information. This transaction takes several

3

phases (i.e., Connect, Request, Response, Close) to
complete. In addition, the requested information from the
server is most commonly an HTTP file which references
graphical images, audio files, and/or additional text. The
5   browser will then automatically request the additional
information (unless the user has manually selected to
filter out certain high bandwidth file requests from being
made). It is not uncommon for an initial request of a web
page to trigger 20 more requests which are required to
10   complete the transaction and build the image seen through
the browser.

THE PRESENT INVENTION

Although the implementation of asymmetrical internet
connections has made internet interaction more efficient,
15   there is yet more compression that can occur to make the
upstream channel even more compact.

In this invention, the upstream traffic between the
Client and the Upstream Gateway is compressed and reduced
with the bulk of the internet communication performed at
20   the Upstream and (in some cases) Downstream Gateway. This
allows more users to be placed on multiple access upstream
channels and/or reduce the user cost on a charge-by-the-
byte upstream channel. This invention exploits the
architecture of the Indirect Network Connection that is

4

utilized in an asymmetrical internet connection as well as the characteristic that the upstream channel from the Client to the Upstream Gateway is a point-to-point connection.

This invention utilizes four types of compression to reduce upstream channel traffic so that the ratio of downstream to upstream channel traffic is increased significantly beyond the 20:1 ratio that is now experienced:

(1) Format compression of request packets.

The typical browser request contains ASCII symbols with characters and patterns that are commonly repeated. For example, in the request:

http://www.cnn.com/SPORTS/FOOTBALL/college/96/news.usarate.html

"/" and "." are the most common characters. Patterns like "http://", "www", ".com", and ".html" are also common. A simple statistical compression or linear predication scheme could reduce the number of bytes required for a request and would be appropriate for this real-time compression application.

Software running at the Client site could operate beneath the browser to compress the ASCII text requests prior to their TCP/IP formatting and the full, uncompressed, request would be regenerated at the Upstream Gateway.

5

(2)   <u>Perform TCP and IP packetizing at the Upstream Gateway</u>
       <u>rather than at the Client:</u>

       Since the upstream channel between the Client and
the Upstream Gateway is a point-to-point link, the
packet overhead of TCP and IP are not necessarily
required. Software running at the Client site would
operate beneath the browser to bypass the TCP/IP
packet formation after the initial connection is
established. The Upstream Gateway would then take the
requests and add the TCP/IP overhead prior to
transmission to the Internet.

(3)   <u>Keep a User Profile at the Upstream Gateway:</u>

       A large database could be implemented at the
Upstream Gateway which stores a Client's previous data
requests. Software running at the Client site would
operate beneath the browser to send a short address
reference message to the Upstream Gateway to indicate
a previously requested site instead of sending an
entire request sequence. The Upstream Gateway, in
turn, would look up the short address reference in the
database for that Client in order to send out the
complete data request to the Internet. This could be
implemented with a "Go back "N" requests" scheme or a
Bookmark scheme similar to those on a typical browser.

6

(4) Automatically generate additional HTTP requests:

The fact that a single web page likely requires multiple data requests from the Client is the largest contributor to upstream data traffic. Instead of the Client system sending the automatic subsequent requests for additional data (referenced by the initial HTTP file delivered from the Server), the gateways could be programmed to receive the first request response from a server and automatically make all of the subsequent file requests for the Client. Each of the requested files would then be downloaded to the Client on the downstream channel as they are received. This makes it possible to have only one data request to be sent by the Client on the upstream channel (instead of a multitude) to assemble a web page.

In addition to the higher HTTP level compression, the gateways could also handle the lower HTTP level of connection, request, response, and close messages with the Internet so that the Client would not be required to send all of these separate messages on the upstream channel.

The Upstream and Downstream Gateways could operate in concert for this service, or they could

7

operate independently with one gateway or the other managing the proxy requests.

Software running at the Client site would operate beneath the browser to manage these compression functions. In addition, should the Client wish to filter out the request of certain files (e.g. to speed up their internet session or to reduce downstream bandwidth), the compression management software could send a message to the gateways to indicate this and only the desired files would be automatically requested by the gateways.

BRIEF DESCRIPTION OF THE DRAWINGS

The above and other objects, advantages and features of the invention will become more clear when considered with the following description and accompanying drawings, wherein:

Figure 1 (prior art) is a diagrammatic illustration of a typical indirect internet connection,

Figure 2 is a block diagram of the client processor for internet browsing with a compressed upstream channel,

Figure 3 is a block diagram illustrating a first compression technique incorporated in the invention,

8

Figure 4 is a block diagram of a second compression technique incorporating the invention,

Figure 5 is a block diagram illustrating a third compression technique incorporated in the invention, and

Figure 6 is a block diagram illustrating a fourth compression technique incorporated in the invention.

DETAILED DESCRIPTION OF THE INVENTION

Figure 1 illustrates all the basic elements of an Indirect Internet Connection where the upstream and downstream data paths are asymmetric and different in their implementation. A Client (i.e., user) (1) requests information from a Server 7 (e.g., home page) somewhere on the internet. It sends the request on the upstream channel (2) to the Upstream Gateway. The Upstream Gateway forwards the request onto the Internet (5) over a higher speed link (4) and it eventually arrives at the addressed server (7). Server 7 responds to the request by addressing the Downstream Gateway (10) and the Client (1). The response arrives at the Downstream Gateway 10 from the Internet (5) over a high speed link (9) and the Downstream Gateway forwards the request to the Client over the Downstream Channel (11). (In most internet service configurations, the two gateways are typically co-located (e.g., dial-up

9

ISP, cable headend; etc.), but it is also possible for the
two gateways (and associated upstream and downstream paths)
to be separate and independent (e.g., satellite broadcast
downstream, telephone line upstream).)

5   Figure 2 shows an exploded view of the Client
Processor CP and the relevant elements that would be needed
for internet web browsing with a compressed upstream
channel (for simplicity, only the upstream path is shown)
according to the invention. The User (11) initiates an

10   information request to the Internet Browser (13) through
the User Interface (12) (e.g., a query to a web page). The
Internet Browser (13) sends a request message to the
appropriate drivers (e.g. TCP/IP, PPP, SLP) of the
operating system (e.g. Windows 95, Solaris, MacOS) where

15   the request message is formatted for transportation to and
on the internet. For the Internet Drivers (14), the Custom
Compression Application (15) (the subject of this
invention) intercepts the request message and performs the
appropriate compression techniques (explained below) prior

20   to delivery to the H/W Interface Driver (16), Communication
Interface (17) and the Communication Path (18) for
transmission to the Upstream Gateway (23) (Figure 3), where
the request message decompression takes place.

         Figure 3 shows a description of the first compression

25 .   technique where the ASCII message format is compressed. A

10

software application will run on the Client (20), "below"
the internet browser, so that all messages sent from the
browser will pass through the compression software and the
ASCII text strings will be reduced by efficiently coding

5   the repeating or redundant characters real-time or near
real-time (using a simple statistical compression or linear
predication scheme), then passed on to the drivers that
transmit the message on the Communication Channel (12) to
the Upstream Gateway (23). The complementary software

10   running at the Upstream Gateway (23) receives the messages
from the Communication Channel (22) and reconstructs the
full ASCII text string prior to delivery to transmission on
the Communication Channel (24) to the Internet.

Figure 4 shows a description of the second compression

15   technique where the TCP and IP packetizing and overhead is
added at the Upstream Gateway (23'). A software
application will run on the Client (20'), "below" the
internet browser, so that all messages sent from the
browser will be intercepted by the compression software and

20   the unnecessary TCP/IP overhead are not added. The
complementary software running at the Upstream Gateway
(23') receives the messages from the Communication Channel
(22) and reads (then removes) the client ID so that the
appropriate TCP/IP overhead can then be added prior to

11

delivery to transmission on the Communication Channel (24) to the Internet.

Figure 5 shows a description of the third compression technique where a short web address reference message is sent to the Upstream Gateway (23''). A software application will run on the Client (20''), "below" the internet browser, which follows the user's web browsing and notes frequently visited sites and also bookmarks that the user enters. The software can then exchange information with the Upstream Gateway (23'') so that a web site address is stored in a user profile database, and referenced by the Client software by a short coded reference message from the Client (20'') to the Upstream Gateway (23'') instead of the complete web site address. In addition, the user could manually set up a number of sites that are to be referenced using this shortcut.

Figure 6 illustrates the fourth compression technique where the Upstream Gateway (23''') requests all additional web page graphics, images, etc. automatically. A software application will run on the Client (20'''), "below" the internet browser, so that the browser is inhibited from automatically requesting any subsequent graphics or image files that may be referenced in the initial web page HTML file response. Instead, when a web page is addressed for delivery by the Client (20'''), the Upstream Gateway

12

(23''') will receive a copy of the initial web page HTML file and make all of the subsequent file requests for the Client (20''') thereby acting as an internet proxy for the Client (20'''). Should the user not wish for the subsequent files to be requested (e.g. in order to save downstream bandwidth), a setting could be sent from the Client (20''') to the Upstream Gateway (23''') which would set up the Upstream Gateway software to not make the requests.

FEATURES OF THE INVENTION

The upstream channel traffic from a Client can be significantly reduced, increasing the number of users on a multiple access upstream channel or reducing the required bandwidth for a single user. This is accomplished by:

1. Compressing the data between the Client and Upstream Gateway.

2. Having the Upstream Gateway perform protocol overhead functions.

3. Keeping a user profile at the Upstream Gateway.

4. Automatically generating new request at the gateway for the client.

One or more of these techniques may be used to "compress" data traffic on the upstream request channel.

13

While the invention has been described in relation to preferred embodiments of the invention, it will be appreciated that other embodiments, adaptations and modifications of the invention will be apparent to those skilled in the art.

14

WHAT IS CLAIMED IS:

1.   In   an   interactive   internet   information
distribution system in which user stations are connected to
the internet by an internet server through an upstream
gateway for transmission of information requests to said
internet server and downstream gateways for transmission of
information from the internet server to the user stations,
the   improvement   wherein   each   user   station   includes   A
processor including:

(a)   means for interfacing with the user,

(b)   an   internet   browser   module,   connected   to   said
interfacing means,

(c)   internet driver means connected to said internet
browser means,

(d)   compression   module   connected   to   said   internet
driver means for reducing the upstream traffic whereby the
ratio   of   downstream   to   upstream   channel   traffic   is
increased significantly beyond a 20:1 ratio.

2.   The information distribution system defined in
Claim 1 wherein said upstream gateway is adapted to perform
protocol overhead functions.

15

3.   The information distribution system defined in Claim 1 wherein said upstream gateway includes means for storing a profile of the users connected thereto.

4.   The information distribution system defined in
5   Claim 2 wherein said upstream gateway includes means for storing a profile of the users connected thereto.

5.   The information distribution system defined in Claim 1 wherein said upstream gateway includes means for automatically generating a new data request for a given
10   user.

6.   The information distribution system defined in Claim 2 wherein said upstream gateway includes means for automatically generating a new data request for a given user.

15   7.   The information distribution system defined in Claim 3 wherein said upstream gateway includes means for automatically generating a new data request for a given user.
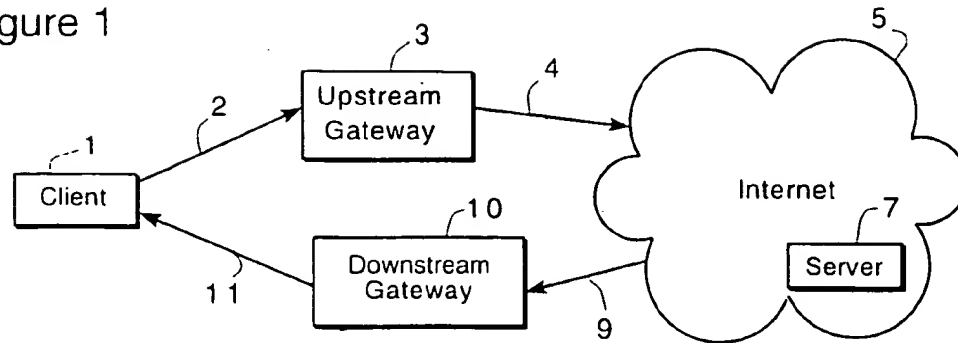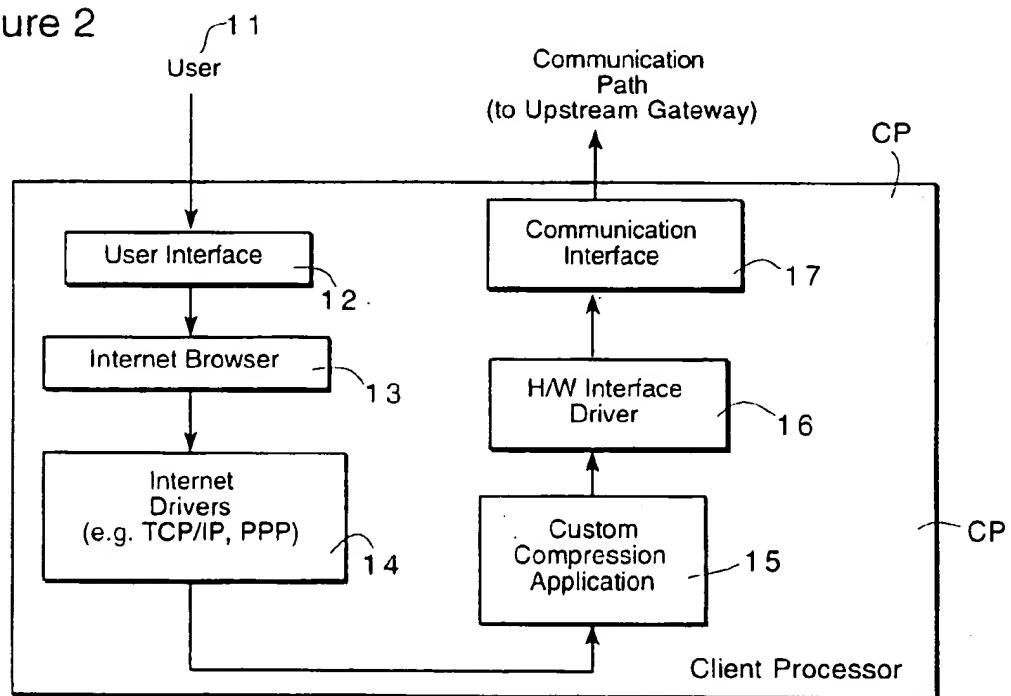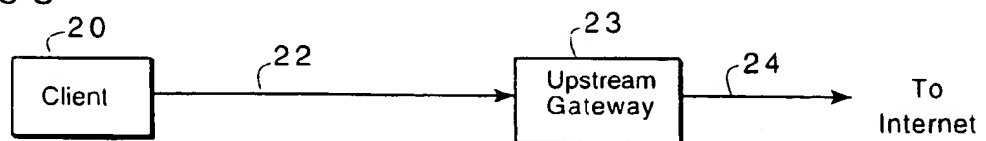
Figure 1

Client — 1

2 → Upstream Gateway — 3

4 → Internet — 5

Server — 7

9 → Downstream Gateway — 10

11 →

Figure 2

User — 11

Communication Path
(to Upstream Gateway)

CP

User Interface — 12

Communication Interface — 17

Internet Browser — 13

H/W Interface Driver — 16

Internet Drivers
(e.g. TCP/IP, PPP) — 14

Custom Compression Application — 15

CP

Client Processor

Figure 3

Client — 20

22 →

Upstream Gateway — 23

24 → To Internet

Figure 4


```
   ┌─20'                    ┌─23'
   ┌─────────┐      22      ┌───────────┐      24'      To
   │ Client  │──────────────│ Upstream  │──────────────→ Internet
   │         │              │ Gateway   │
   └─────────┘              └───────────┘
```

Figure 5


```
   ┌─20"                    ┌─23"
   ┌─────────┐      22      ┌───────────┐      24      To
   │ Client  │──────────────│ Upstream  │──────────────→ Internet
   │         │              │ Gateway   │
   └─────────┘              └───────────┘
```

Figure 6


```
     20'''                    23'''
   ┌─────────┐      22      ┌───────────┐      24      To
   │ Client  │──────────────│ Upstream  │──────────────→ Internet
   │         │              │ Gateway   │
   └─────────┘              └───────────┘
```
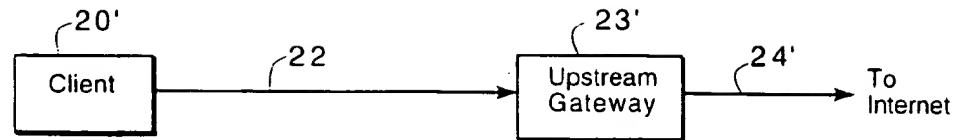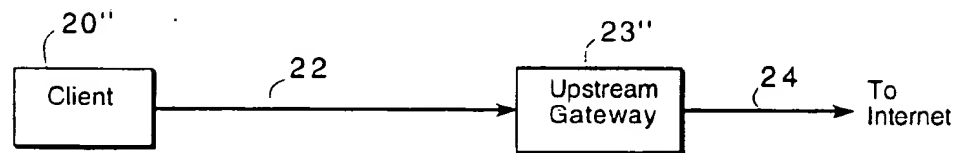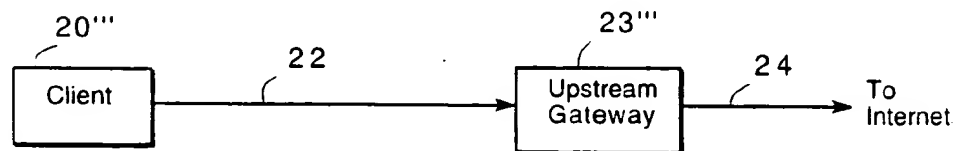
INTERNATIONAL SEARCH REPORT

A.	CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G06F 13/42, 15/173
US CL :395/200.48, 200.77

According to International Patent Classification (IPC) or to both national classification and IPC

B.	FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 395/200.47, 200.48, 200.49, 200.69, 200.77; 370/352

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

NONE

C.	DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, E	US 5,742,773 A (BLOMFIELD-BROWN et al.) 21 April 1998, see Abstract; Fig. 2-3, 5; C1:14-28; C2:29-59; C4:32-C5:45; C6:35-C9:30	1-7
A, P	US 5,673,392 A (NAKASHIMA et al.) 30 September 1997, see Abstract; Fig. 1-3; C5:1-C6:60	1-7
A	US 5,557,749 A (NORRIS) 17 September 1996, see Abstract; Fig. 6; C5:53-C6:30	1-7
X	US 5,555,377 A (CHRISTENSEN et al.) 10 September 1996, see Abstract; Fig. 1-2, 5; C3:20-34; C4:47-67	1-7

[X] Further documents are listed in the continuation of Box C.   [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
		"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 MAY 1998	0 4 AUG 1998

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer MARK H. RINEHART
Facsimile No. (703) 305-3230	Telephone No. (703) 305-9700

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,534,913 A (MAJETI et al.)09 July 1996, see Abstract; Fig. 1; C3:4-C5:62	1-7